International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Finger Reader for Visually Impaired People Using Optical Character Recognition (OCR) and Machine Learning Technique

Mr. Ajith G L [1], Mr.Musheer Ahmed [2].

Mr. Ajith G L Assistant Professor, Dept. of MCA PESITM, Shimoga, Karnataka

Mr.Musheer Ahmed Assistant Professor, Dept. of MCA PESITM, Shimoga, Karnataka

Abstract

Finger Reader is a wearable gadget that uses machine learning algorithms and a camera positioned on the finger to read text for those who are blind or visually impaired. The device takes text pictures and recognizes letters, words, and phrases using convolutional neural networks (CNNs). The World Health Organization estimates that 253 million people globally are blind or visually impaired. From the perspective of the visually impaired, there are a number of inconveniences that slowdown their day-to-day activities. To enable them to live in the modern world without difficulty, it is imperative that significant steps be taken using the technology currently in use. The system utilizes a combination of optical character recognition (OCR) and natural language processing (NLP) techniques to provide accurate and context-aware text recognition.

The Device also includes a tactile feedback system, providing users with a physical representation of the text. Experimental results demonstrate high accuracy rates in text recognition, making Finger Reader a promising assistive technology for visually impaired individuals to access written information independently.

Keywords: Optical character Recognition (OCR), Natural Language Processing (NLP), Convolutional neural network(CNN), Machine learning(ML), Finger Reader.

1. Introduction

People who are blind or visually challenged may read printed words or texts thanks to the Finger Reader. For blind individuals, accessing textual content on a mobile device might be challenging. Essentially, it is an instrument wear on the finger that includes a tiny digital camera for response. They can move back and forth, faster and slower. We'll show you the simplest way to scan text so you can read language or a line of text while receiving real-time feedback—both touch and sound.

The system is light enough to operate without the usage of eyes. When people having impaired vision want to read nearby text, such as a restaurant menu card, they simply point their finger at it, and the gadget reads the words in the correct order. With a finger reader, sometimes called a text reader or finger-worn scanner, a person with visual impairments may perform nearly all tasks that those who are blind can perform.

This device can be used to read text in front of them in any language and typeface, helping them to learn about the world surrounding them. It enables students to study any book, which is considered a fundamental part of everyone's life. The majority of blind people are unaware of the reading system, which they utilize for learning purposes. This solution reduces the need to convert every text into braille by sending it straight as audio. This method exposes the user to various kind of information than simply reading items.

2. Literature Survey

The visual reading aids now available for the blind have a number of shortcomings. Braille, a tactile method of raised dots for letters, is vital for the blind and visually impaired, however not all publications are accessible in this format. As such, the amount of material available to the blind is restricted. Current technologies suffer from problems with accuracy, portability, efficiency, and attentiveness.[1]

The paper proposes a reading aid for the blind that aims to faithfully replicate common printed materials. A dense array of vibrating pins that may be individually actuated is used to accomplish this. The tactile picture is felt by the user by placing their finger on a plate with holes that match the pin array. The proposed technique positions an array of photocells, each linked to piezoelectric reeds that regulate the pins producing the tactile picture, to align the image of standard printed text. The study assesses the setup's practicality and looks at the links between different design aspects. [2]

The paper introduces the concept design of HandSight, an advanced vision- augmented touch system aimed at sensing and providing non-tactile information about the environment to assist with Activities of daily living (ADLs). While reading printed text is the main focus, expanding its applications to different ADL contexts is the ultimate goal. The report presents the current deployment state of HandSight, its development ambitions, and preliminary study results on the effectiveness of finger-based text scanning.[3]

3. Proposed Methodology

The main goals are to translate user-pointing photos into text files so that visually impaired individuals may convert them into sounds. As part of the setup, a high-resolution camera is placed on the user's hands. There is a button on the fingers that starts the picture-taking procedure.

A Raspberry Pi uses a number of algorithms to

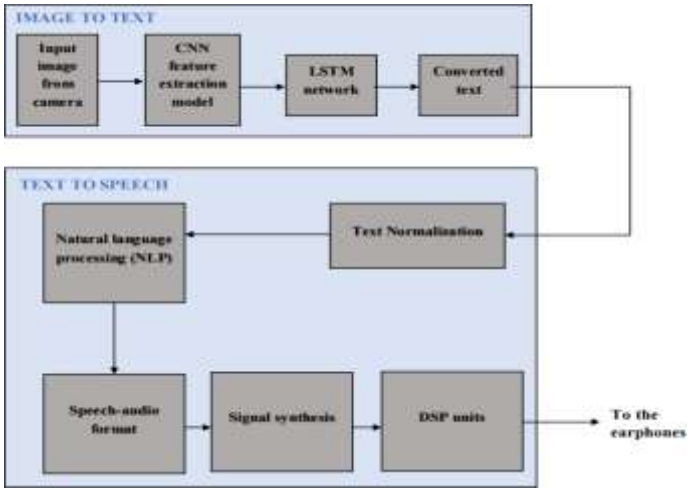

Fig 3.1: System Design

The image is taken by the camera and saved in the Raspberry Pi's external memory when the user hits the button while pointing their finger at the text they want to read. A fully convolutional neural network (FCN), a neural network-based model, is used to extract text from the acquired image. This FCN provides dense per-pixel predictions of the text's geometry by directly detecting text occurrences from the image.

Using FCN in this case has the benefit of streamlining the procedure by doing away with intermediary procedures like word partitioning, candidate proposal, and text region construction. The collected text is subjected to a few minor preprocessing processes, such as thresholding and non-maximum suppression (NMS), following FCN processing.

3.1 Image to text

A finger reader often consists of a camera or scanner that takes text images, as well as OCR (Optical Character Recognition) software that converts these images into machine-readable text. When a user places their finger on the device and drags it across printed text, the camera takes a series of photos. The OCR program then compares the photos to a database of known characters. To recognize extract text from the captured image and then turns it into audio that is played back over headphones.

Considering the modifications, finger readers continue to provide a convenient and lightweight means of accessing printed content for anyone with vision impairments or reading difficulties. These devices boost the convenience and flexibility of reading books, papers, and other materials by converting written content into digital format.

Details on various steps:

3.1.1 Input image from camera:

In a finger reader, the procedure starts with capturing an input image with a camera or scanner located at the device's bottom. The user places their finger on top of the camera to begin capturing images of the text. When the user swipes their finger across text, the camera takes pictures. These photos are often captured at high resolution to ensure clarity and accuracy when using OCR (Optical Character Recognition) software to detect characters within the text. The camera in a finger reader can use CMOS (Complementary Metal-Oxide Semiconductor) or CCD (Charge-Coupled Device) sensors. These sensors transform light reflected from the text to digital format, later it will processed by OCR software.

3.1.2 CNN feature extraction model:

A Convolutional Neural Network (CNN) is a key unit in a finger reader's image processing activities, particularly feature extraction for OCR (Optical Character Recognition). The CNN model analyzes and extracts information from images taken by the device's camera. Convolutional layers, which utilize filters to find patterns in input images, pooling layers, which reduce feature maps to smaller samples, and fully connected layers, which interpret extracted features for tasks like classification or. This complete method ensures

recognition, are some of the layers that make up CNNs. In finger reader, the CNN feature extraction model is trained on a heterogeneous dataset containing multiple types of text, typefaces, and size. The model learns to identify various characters and their attributes in text images. Once trained, the CNN evaluates the input images taken by the fingerprint reader's camera, extracting relevant information characters in photos, OCR software uses a variety of approaches including pattern recognition, feature extraction, and machine learning algorithms.

3.1.3 LSTM network:

In a finger reader, after the CNN (Convolutional Neural Network) extracts digital text from collected pictures, an LSTM (Long Short-Term Memory) network converts this text into spoken output using a synthetic voice. The LSTM network may be trained using a variety of languages and voices, allowing users to customize the audio output to their preferences. The usage of an LSTM network improves audio output accuracy and naturalness by successfully simulating the text's long-term dependencies. This capacity allows the network to produce fluent and understandable speech that mimics actual human speech patterns.

3.1.4 Converted text:

In a finger reader, text conversion involves using a camera to capture an image of printed text, which is then enhanced for clarity using techniques such as filtering and contrast correction. The augmented image is subsequently processed using Optical Character Recognition (OCR) software, which segments the characters, extracts characteristics, and uses machine learning techniques to recognize and transform them into digital text. To improve accuracy and readability, post-processing includes spell-checking, grammatical correction, and formatting changes to the transformed text.

blind people or reading challenges to properly access printed materials via digital text output.

3.2 Text to speech

Text-to-speech (TTS) technology is essential to a finger reader's operation. After Optical Character Recognition (OCR) software translates text images into digital text, TTS software takes over, converting the text into spoken words with a computer-generated voice. TTS software scans digital text and applies pronunciation and intonation rules to produce a natural-sounding voice. It also understands punctuation to ensure that the spoken output contains suitable pauses and inflections. Modern TTS technology has improved dramatically, resulting in voices that are both natural and easy to understand. TTS software often delivers audio output through speakers or headphones in a finger reader configuration. Users may update settings like loudness to ensure that audio is clear and comprehensible.

3.3 Text Normalization:

Text normalization is important in finger readers because it standardizes text into a format that computers can effectively interpret, boosting accuracy and the reading experience for users. Text normalization involves standardizing text representation by translating variances like letter cases, punctuation marks, and special characters into a consistent format. It considers abbreviations, acronyms, and shorthand forms using Natural Language Processing (NLP) algorithms. Finger readers use text normalization to manage multiple languages and writing systems, increasing recognition accuracy and enhancing reading skills for visual impairments or reading challenges.

CNN (Convolutional Neural Network):

3.4 Natural Language Processing (NLP):

Natural Language Processing (NLP) improves the ability of finger readers by allowing them to analyze, translate, and synthesize text, which benefits those with visual impairments or reading issues. These devices' NLP algorithms can recognize text on a variety of surfaces and transform it to speech or Braille for accessibility. This feature enables users to interact with printed materials autonomously.

4.Technology Used:

Fig 4.1: System Architecture

To use a finger reader, position it on the index finger, line it up with the text, take a picture, use OCR and image processing algorithms to process it, turn it into digital text, have a text-to-speech engine read it out loud or show it on a device that is connected, and repeat as necessary. Certain finger readers can link to external devices or have built-in storage for future use. More sophisticated devices may use machine learning algorithms to identify different printed text styles, languages, and typefaces

A Convolutional-Neural-Network (CNN) is a deep learning algorithm primarily utilized for image recognition and classification. It consists of three main types of layers:

Convolutional Layer: The basic component of a CNN is the convolutional layer, which detects local patterns in input data such as edges, textures, and forms. It performs a mathematical action known as convolution by sliding filters (kernels) across the input image. At each place, the filter multiplies its values by the matching input values and adds them together to produce a single value in the final feature map.

Pooling Layer: The pooling layer decreases the spatial dimensions (width and height) of the feature maps created by the convolutional layer while maintaining the most critical data. Max-pooling and average-pooling are two common pooling techniques. Max pooling selects the maximum value inside a certain window, whereas average pooling finds the average value within that frame.

Fully Connected Layer: The fully-connected layer, also known as the dense layer, is comparable to a standard neural network layer in that each neuron is linked to every neuron in the preceding layer. This layer takes as input a simplified version of the previous layer's feature maps. Each input value is multiplied by a weight, and the results are summed with a bias term before being routed through an activation function.

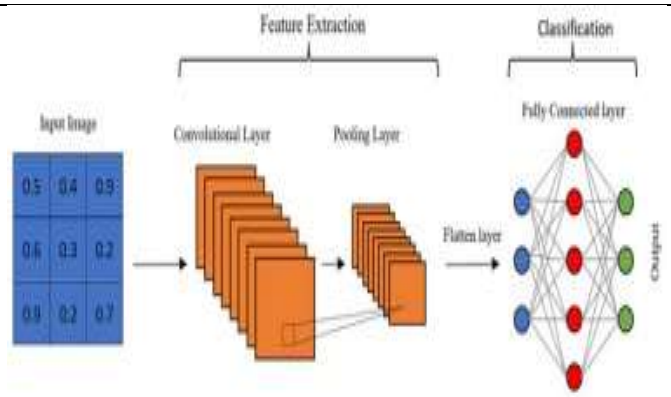

Fig 4.2: CNN Architecture

5.Results and Discussion

Fig 5.1: Reading Text on E-Book

Fig 5.2 : Reading Text on a Book

technologies: reading blocks of text at a time.

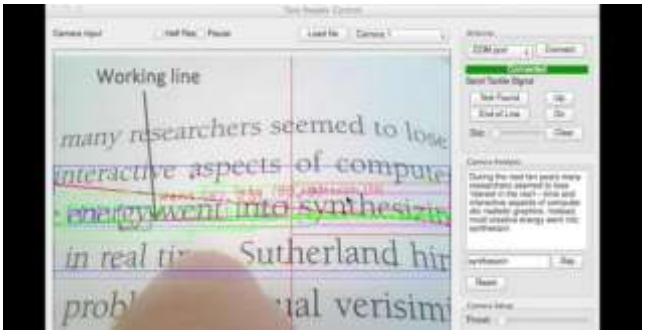

Fig 5.3: Text Recognition in a TTS Software

Fig 5.4 : Text Recognition While finger sliding on the text

6.Conclusion

Finger readers are a type of assistive technology that gives sensory or audible feedback to those with visual impairments or reading challenges, allowing them to read more easily and independently. Finger readers have several advantages, including portability, customization, versatility, non-intrusiveness, and affordability. However, they may have certain drawbacks, including poor tactile feedback, speed and accuracy, availability, language support, and compatibility. A innovative approach for blind text reading that uses a local-sequential scan to provide continuous feedback and non-linear text scanning. Motivated by focus group sessions with blind participants, our strategy presents a solution to a

Our solution incorporates a text tracking algorithm that extracts words from a close-up camera view, which is coupled with a finger-worn device. A technical accuracy examination revealed that the local-sequential scan algorithm is reliable. Two qualitative investigations with blind individuals provided vital insights into the growing field of finger-worn reading aids. Overall, finger readers are a useful, effective, and accessible tool for people with visual impairments or reading challenges, but their efficacy and suitability should be assessed on an individual basis with the assistance of a skilled professional or assistive technology specialist.

References

[1] Pooja Deole, Shruti Kulkarni, "FINGER READER: A WEARABLE DEVICE FOR THE VISUALLY IMPAIRED", International Journal of Infinite Innovations in Technology ISSN:2278-9057|IJIIT|Volume-IV|Issue-II|2015-2016 October

[2] J. G. LmvILL," A Direct Translation Reading Aid for the Blind", Proceedings of the IEEE vol. 54, no. 1 January, 1966.

[3] Bigham, J. P., Jayant, C., Ji, H., Little, G., Miller, A., Miller, R. C., Miller, R., Tatarowicz, A., White, B., White, S., and Yeh, T. Vizwiz: Nearly real-time answers to visual questions. In Proc. of UIST, ACM , 333–342

[4] Andrea R. Kennel, "Audiograf: A Diagram-Reader For The Blind", ASSETS '96, Vancouver, British Columbia, Canada 0 1996 ACM 0-89791-176-6/96.

[5] Singh, S., et al. "Development of a Wearable Device for Visually Impaired People." Proceedings of the 4th International Conference on Advances in Computing, Communication & Control.

common shortcoming of most existing

[6] Talukdar, S., et al. "A Smart Glove for Blind Navigation and Reading." Proceedings of the 6th International Conference on Computer and Communication Engineering.

[7] Chen, C., et al. "A Finger Reading System for Blind and Visually Impaired People." Proceedings of the 13th International Conference on Computer and Information Technology.

[8] Pal, D., et al. "Finger Reader: A Wearable Device to Support Text Reading on the Go." Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility.

BIOGRAPHIES

Mr. Ajith G L,
Assistant Professor,
Dept of MCA,
PES Institute of Technology and Management, Shimoga.

Mr. Musheer Ahmed,
Assistant Professor,
Dept of MCA,
PES Institute of Technology and Management, Shimoga.